# Modeling nonstationary noise in pulsar timing array data analysis

Mikel Falxa,[1,2,*] J. Antoniadis,[3,4] D. J. Champion,[4] I. Cognard,[1,5] G. Desvignes,[4] L. Guillemot,[1,5] H. Hu,[4] G. Janssen,[6,7] J. Jawor,[4] R. Karuppusamy,[4] M. J. Keith,[8] M. Kramer,[4] K. Lackeos,[4] K. Liu,[9,4] J. W. McKee,[10,11] D. Perrodin,[12] S. A. Sanidas,[8] G. M. Shaifullah,[2,13,12] and G. Theureau[1,5,14]

[1]*Laboratoire de Physique et Chimie de l'Environnement et de l'Espace, Université d'Orléans/CNRS, 45071 Orléans Cedex 02, France*
[2]*Dipartimento di Fisica "G. Occhialini," Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*
[3]*FORTH Institute of Astrophysics, North Plastira 100, 70013 Heraklion, Greece*
[4]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*
[5]*Observatoire Radioastronomique de Nançay, Observatoire de Paris, Université PSL, Université d'Orléans, CNRS, 18330 Nançay, France*
[6]*ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands*
[7]*Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands*
[8]*Department of Physics and Astronomy, Jodrell Bank Centre for Astrophysics, University of Manchester, Manchester M13 9PL, United Kingdom*
[9]*Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China*
[10]*E. A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Kingston-upon-Hull HU6 7RX, United Kingdom*
[11]*Centre of Excellence for Data Science, Artificial Intelligence and Modelling (DAIM), University of Hull, Cottingham Road, Kingston-upon-Hull HU6 7RX, United Kingdom*
[12]*INAF–Osservatorio Astronomico di Cagliari, via della Scienza 5, 09047 Selargius (CA), Italy*
[13]*INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*
[14]*Laboratoire Univers et Théories LUTh, Observatoire de Paris, Université PSL, CNRS, Université de Paris, 92190 Meudon, France*

(Received 11 February 2024; revised 20 April 2024; accepted 6 May 2024; published 4 June 2024)

Pulsar timing array (PTA) collaborations recently reported evidence for the presence of a gravitational wave background (GWB) in their datasets. The main candidate that is expected to produce such a GWB is the population of supermassive black hole binaries. Some analyses showed that the recovered signal may exhibit time-dependent properties, i.e., nonstationarity. In this paper, we propose an approximated nonstationary Gaussian process model obtained from the perturbation of stationary processes. The presented method is applied to the second data release of the European Pulsar Timing Array to search for nonstationary features in the GWB. We analyzed the data in different time slices and showed that the inferred properties of the GWB evolve with time. We find no evidence for such nonstationary behavior and the Bayes factor in favor of the latter is $\mathcal{B}_S^{\mathrm{NS}} = 1.5$. We argue that the evolution of the GWB properties most likely comes from the improvement of the observation cadence with time and better characterization of the noise of individual pulsars. Such nonstationary GWB could also be produced by the leakage of nonstationary features in the noise of individual pulsars or by the presence of an eccentric single source.

DOI: 10.1103/PhysRevD.109.123010

## I. INTRODUCTION

Supermassive black holes (SMBHs) are incredibly massive, over a million times heavier than the Sun, and are typically located at the centers of galaxies. When these black holes come together, they form binary systems that emit strong gravitational waves as they approach each other [1]. The combined effect of many such binaries produces a continuous and stationary noise signal in the nanohertz frequency band [2], known as the gravitational wave background (GWB). Pulsar timing array (PTA) collaborations search for the induced effect of this GWB in the timing measurements of millisecond pulsars. They specifically target the characteristic spatial correlations it would

*mikel.falxa@unimib.it

 

exhibit between pairs of pulsars, known as the Hellings-Downs (HD) correlation pattern [3]. Recent studies have shown strong evidence for the presence of such a signal [4–6]. However, it does not exactly match what was expected for a population of supermassive black hole binaries (SMBHBs) in circular orbits. In particular, it was observed that the inferred GWB properties differed when datasets of different lengths were analyzed. Several hypotheses have been proposed to explain this discrepancy, one of which is the possible nonstationarity of the GWB.

The European Pulsar Timing Array (EPTA) Collaboration has released its second dataset, combining the timing data of 25 pulsars over a span of up to 24.7 years of observation. Several radio telescopes were used to collect the data: the Effelsberg Radio Telescope in Germany, Lovell Telescope in the UK, Nançay Radio Telescope in France, Westerbork Synthesis Radio Telescope in the Netherlands, Sardinia Telescope in Italy, and Large European Array for Pulsars. In their analysis [6], the EPTA searches for the presence of a GWB in the dataset by evaluating the Bayes factor, comparing common HD correlated noise against common uncorrelated noise denoted as $\mathcal{B}^{\mathrm{HD}}_{\mathrm{CURN}}$. The significance and properties of the GWB vary notably across the dataset. When analyzing the entire dataset, the Bayes factor is only $\mathcal{B}^{\mathrm{HD}}_{\mathrm{CURN}} = 4$, but it is shown that, using only the latest 10.33 years of data, the Bayes factor rises to $\mathcal{B}^{\mathrm{HD}}_{\mathrm{CURN}} = 60$. The justification for trimming the dataset is the evolution of data quality. Legacy data have a worse cadence of observation and fewer multifrequency band observations [7]. These two qualities are crucial for performing a good characterization of all sources of noise present in the timing data, essential for an accurate observation of the GWB [8]. Radio observatories are equipped with receiver-backend systems where the receiver collects the raw radio data that goes through a real-time processing stage in the backend systems, providing analyzable data [7]. The latest 10.33 years of data were obtained from new-generation backends with greatly improved observing abilities. However, the evolution of properties of the observed signal was not only reported by the EPTA Collaboration. The Parkes Pulsar Timing Array performed a time-slice analysis where their dataset was analyzed in slices of 3 or 9 years and showed that the amplitude of the signal varied with time [5]. The reason behind this time-evolving property remains unexplained as the GWB is expected to be stationary [1,2]. One advanced hypothesis is nonstationarity of noise present in the pulsars themselves that would contaminate our estimate of the signal. The two dominant sources of time-correlated noise in pulsar timing are the red noise (RN), due to stochastic variation of the pulsar's spin rate, and the dispersion measure (DM) noise, due to the fluctuation of the electron density in the interstellar medium during the path that the pulses take to Earth [8] (this noise has an amplitude proportional to $\nu^{-2}$ with $\nu$ being the incoming photon's frequency). The noise contributions are typically assumed to be stationary. Other potential sources of nonstationarity of astrophysical origin are eccentric SMBHBs. They produce gravitational waves (GWs) with significantly varying frequency content within one orbital period of the binary [9,10]. The presence of one or more eccentric sources could influence a stationary GWB and make it nonstationary.

In this paper, we propose a way to model nonstationary noise. Previous works explored the idea of modeling nonstationary signals in GW analysis using wavelets [11,12]. Here, we develop a model that is a perturbation of the Gaussian process (GP) methods that are commonly used in PTA data analysis [13]. The main motivation behind this model is to evaluate the significance of nonstationary features in noise through the evaluation of the Bayes factor by allowing time dependence of the noise spectrum properties. Strong variations in time result in strong correlations between different frequencies of the spectrum that are not accounted for by ordinary Fourier decomposition [12]. In general, the data are cut into smaller intervals where the noise is considered approximately stationary, and a Fourier transform is applied in each window to monitor the evolution of the frequency content. The assumption of local stationarity presupposes that the properties of the spectrum evolve slowly in the analyzed interval [14].

This paper is organized as follows. In the first section, we will present a nonstationary GP model. The ability of this model to correctly infer parameters is then tested on mock datasets in which nonstationary noise is injected. Finally, we will search for the possible presence of a nonstationary GWB in the EPTA dataset and argue that nearby eccentric SMBH binaries could introduce nonstationary features in the GWB.

## II. MODEL

### A. Stationary Gaussian process

A stationary process is a stochastic process whose properties, typically its mean and variance, remain constant over time. In practice, we often encounter what we refer to as weak-sense stationary processes, where the mean and correlation functions remain invariant under shifts in time, i.e., they are time-translation symmetric [15]. Given the one-sided power spectral density (PSD) $S_0(f)$ of the process, the correlation function $C(t,t')$ can be expressed using the Wiener–Khinchin theorem integral,

$$C(t,t') = \int_0^\infty df S_0(f) \cos(2\pi f(t-t')), \qquad (2.1)$$

which is a function of the time difference $t - t'$.

Alternatively, we can define a zero-mean stationary Gaussian process expressed as a finite Fourier sum with normally distributed coefficients [13,16],

$$n_0(t) = \sum_{i=1}^{N} X_i \sin(2\pi f_i t) + Y_i \cos(2\pi f_i t) = \sum_{i=1}^{N} \vec{\omega}_i \cdot \vec{\phi}_i(t), \tag{2.2}$$

where $\vec{\phi}_i(t) = [\sin(2\pi f_i t), \cos(2\pi f_i t)]$ and $\vec{\omega}_i = [X_i, Y_i]$ with $X_i, Y_i \sim \mathcal{N}(0, \sigma_i^2)$.

The basis $\vec{\phi}_i(t)$ of cosine and sine functions is evaluated at a discrete set of frequencies $f_i$. For stationary signals, the random weights $\vec{\omega}_i$ are uncorrelated because the underlying Fourier basis of the process is orthogonal. Thus, using $\langle \vec{\omega}_i, \vec{\omega}_j \rangle = \langle \omega^2 \rangle ij = \sigma_i^2 I_2 \delta ij$ (where $\langle a, b \rangle$ denotes the ensemble averaged inner product between $a$ and $b$, and $I_2$ is the identity matrix of rank 2), we have

$$C(t,t') = \langle n(t), n(t') \rangle = \sum_{ij} \vec{\phi}_i^{\top}(t) \langle \omega^2 \rangle_{ij} \vec{\phi}_j(t') \equiv \sum_i \sigma_i^2 \cos(2\pi f_i (t - t')), \tag{2.3}$$

which is an approximation of the integral in Eq. (2.1) for $\sigma_i^2 = S_0(f_i)\Delta f_i$ with $\Delta f_i = f_{i+1} - f_i$ [17].

### B. Nonstationary Gaussian process

By definition, a nonstationary process exhibits spectral properties that vary with time [18,19]. Let us consider a time-dependent evolutionary power spectral density defined as

$$S(f,t) = S_0(f) \times g^2(f,t), \tag{2.4}$$

where $S_0(f)$ represents a stationary PSD, and $g(f,t)$ is an arbitrary function of time that introduces perturbations to the stationary PSD.

Using this approach, we extend the definition of the zero-mean GP presented in Eq. (2.2) by introducing time-varying coefficients $X_i(t)$ and $Y_i(t)$, where $X_i(t)$, $Y_i(t) \sim \mathcal{N}(0, \sigma_i^2(t))$ and $\sigma_i^2(t) = S_0(f_i) \times g^2(f_i, t)\Delta f_i$. By applying the identity $\mathcal{N}(0, a^2\sigma^2) = a\mathcal{N}(0, \sigma^2)$, the function $g(f,t)$ factors out of the Gaussian weights, yielding

$$n(t) = \sum_{i=1}^{N} X_i(t_0) g(f_i,t) \sin(2\pi f_i t) + Y_i(t_0) g(f_i,t) \times \cos(2\pi f_i t) = \sum_{i=1}^{N} \vec{\omega}_i \cdot \vec{\Phi}_i(t) \tag{2.5}$$

with $\Phi_i(t) = g(f_i,t) \times \phi_i(t)$.

This expression is identical to the stationary case, except the basis $\Phi_i(t)$ can now be modulated over time through the function $g(f_i,t)$. This modulation of the basis gives a description of the nonstationarity of the signal (see Fig. 1). In this context, the relationship $\langle \vec{\omega}_i, \vec{\omega}_j \rangle = \langle \omega^2 \rangle_{ij}$ does not necessarily maintain a diagonal form because the weights are not expected to be uncorrelated. However, we proceed by making an assumption of diagonality. In other words, we assume that the basis $\Phi_i(t)$ is the natural (eigen)basis of the process. In that case, the covariance matrix becomes

$$C(t,t') = \sum_{ij} \vec{\Phi}_i^{\top}(t) \langle \omega^2 \rangle_{ij} \vec{\Phi}_j(t') \simeq \sum_{i}^{N} \sigma_i^2 g(f_i,t) g(f_i,t') \cos(2\pi f_i (t - t')). \tag{2.6}$$

This kernel can be seen as a specific instance of nonstationary generalized spectral kernels (see Refs. [20,21] and the Appendix). It is simplified by considering stationary frequency components $f_i$ accompanied by time-varying weights (a consequence of the assumption that $\langle \vec{\omega}_i, \vec{\omega}_j \rangle = \langle \omega^2 \rangle_{ij} = \sigma_i^2 I_2 \delta_{ij}$ is diagonal). This assumption is reasonable if the evolution of frequencies $f_i$ of the basis $\Phi_i(t)$ is negligible during the considered time span of observation $T$ (i.e., stationary phase). Such drifts in $f_i$ (typically chirplike signals) would produce correlations between consecutive frequencies that are not accounted for in the expression of this kernel. Nevertheless, modeling the

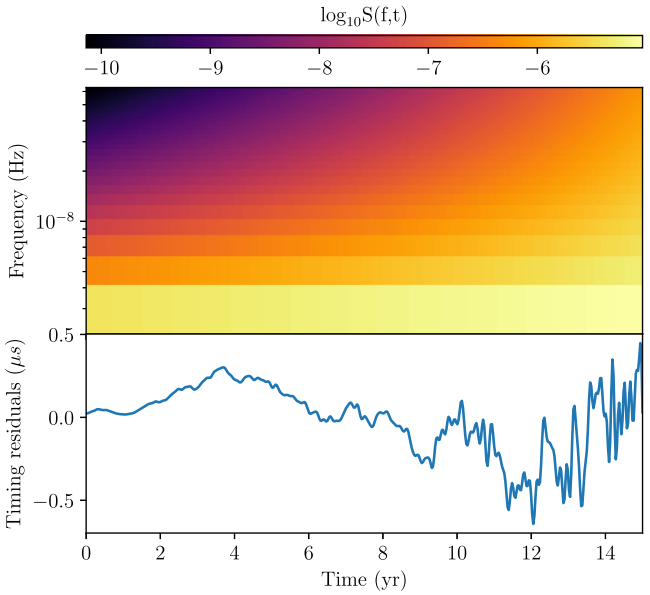


FIG. 1. Example of a simulated nonstationary colored noise for $\log_{10} A(t_0) = -14$ and $\gamma(t_0) = 2$ observed in the timing residuals of a PTA pulsar with 15 years of observation. The noise was simulated using a first-order Taylor expanded power law [see Eq. (4.1)] with $a = 0.8$ and $b = -1.2$, meaning that we have a noise with increasing amplitude and decreasing spectral index (i.e., the noise whitens with time). Top: the evolutionary PSD $S(f,t)$. Bottom: the corresponding time-domain representation.

time-dependent variance of each weight with this expression still captures information about nonstationary behaviors through the function $g(f,t)$.

The previous expression is valid for slowly evolving spectra that can be considered locally stationary. The notion of local stationarity is defined in [14] as when the signal is approximately stationary over some sliding interval $\tau$. In our case, because we want to model slowly evolving spectra, the stationarity timescale $\tau$ should be comparable to the total time of observation $T$. This condition translates to the function $g(f,t)$ as having most of its Fourier spectrum concentrated at frequencies around $1/T$. Essentially, we approximate nonstationary kernels by concentrating the 2D spectral measure $S(f,f')$ along the diagonal where $f=f'$ (see Appendix A 1).

In the upcoming section, we propose conducting simulations to assess the impact of this approximation on parameter recovery. The primary goal is to estimate whether neglecting correlations between Gaussian weights in Eq. (2.6) significantly influences the inference of parameters within a Bayesian framework.

### C. Colored noise

In PTAs, pulsar noise is typically described with a power-law spectrum $S_0(f) \propto Af^{-\gamma}$, where $A$ represents the amplitude and $\gamma$ denotes the spectral index [8,22,23]. Let us consider a time-dependent power spectral density with time-varying amplitude and spectral index

$$S(f,t) = \frac{A^2(t)}{12\pi^2}\left(\frac{f}{f_{yr}}\right)^{-\gamma(t)} f_{yr}^{-3}, \tag{2.7}$$

with $f_{yr}$ the frequency corresponding to one year. Any arbitrary functions of time for $A(t)$ and $\gamma(t)$ can be used as long as we ensure $S(f,t)>0$. We choose to express the amplitude and spectral index as

$$\begin{aligned}\log_{10}A(t) &= \log_{10}A(t_0) + \sum_k a_k P_k(t-t_0),\\ \gamma(t) &= \gamma(t_0) + \sum_k b_k P_k(t-t_0),\end{aligned} \tag{2.8}$$

where $t_0$ denotes the reference time, $a_k$ and $b_k$ are constant coefficients, and $P_k$ are polynomial functions (such as Chebyshev, spline, etc.). We can factor out the stationary PSD $S_0(f) \propto A^2(t_0)f^{-\gamma(t_0)}$ from a time-dependent function $g(f,t)$,

$$\begin{aligned}S(f,t) &= S(f,t_0)\times 10^{2\sum_k a_k P_k(t-t_0)}\left(\frac{f}{f_{yr}}\right)^{-\sum_k b_k P_k(t-t_0)}\\ &= S_0(f)\times g^2(f,t),\end{aligned} \tag{2.9}$$

providing an expression for the function $g(f,t)$ as presented in the previous section. According to Eq. (2.6), we then have

$$\begin{cases} C(t,t') = \sum_i \sigma_i^2 g(f_i,t)g(f_i,t')\cos(2\pi f_i(t-t')) \\ \sigma_i^2 \propto A^2(t_0) f_i^{-\gamma(t_0)}\Delta f_i \\ g(f,t) = 10^{\sum_k a_k P_k(t-t_0)}\left(\frac{f}{f_{yr}}\right)^{-\sum_k \frac{b_k}{2} P_k(t-t_0)}. \end{cases} \tag{2.10}$$

In the next sections of this paper, we will only utilize the first-order term of this expression. We will assume slowly varying properties of the signal.

## III. INJECTION RECOVERY TEST

### A. Likelihood

To perform inference of the parameters, we need to define a likelihood. The GP model previously described assumes that the noise in the data is Gaussian. Hence, for a time series $s(\vec{t})$ of length $N$ measured at times $\vec{t}$, we will consider a Gaussian likelihood of the form

$$\begin{aligned}\mathcal{L}(s|\vec{\theta}) = &\frac{1}{|2\pi C(\vec{\theta})|^{\frac{1}{2}}}\\ &\times \exp\left\{-\frac{1}{2}\sum_i\sum_j s(t_i)C^{-1}(\vec{\theta},t_i,t_j)s(t_j)\right\},\end{aligned} \tag{3.1}$$

where $C^{-1}$ is the inverse of the covariance matrix $C$, $\vec{\theta}$ is the vector of the model parameters, and $|.|$ denotes the determinant.

In the Bayesian framework, we define prior probability distributions $\pi(\vec{\theta})$ for the parameters $\vec{\theta}$ of the model. Combining the prior distributions with the likelihood, we can update our knowledge about $\vec{\theta}$, defining the posterior distribution $p(\vec{\theta}|s)$,

$$p(\vec{\theta}|s) = \mathcal{L}(s|\vec{\theta})\pi(\vec{\theta})/\mathcal{Z}, \tag{3.2}$$

with $\mathcal{Z}$ the so-called "evidence" of the model acting as a normalizing constant.

The key quantity to evaluate the significance of model $A$ against model $B$ in Bayesian analysis is the Bayes factor $\mathcal{B}^A_B$. It is defined as the ratio of the evidences

$$\mathcal{B}^A_B = \frac{\mathcal{Z}_A}{\mathcal{Z}_B}. \tag{3.3}$$

In this paper, the posterior distribution is sampled using Markov chain Monte Carlo (MCMC) sampler PTMCMC [24] or nested sampling library DYNESTY [25].

### B. Time-varying filter

In order to test our model, we want to simulate a colored noise with evolving properties in time using a method that is independent of the one developed in the previous section. Various methods are available to generate such noise, for

example, one can utilize autoregressive models [12]. Here we chose to generate nonstationary data with a time-varying filter that allows one to control the amplitude and spectral index of the noise at each instant [14,18].

We first generate an evenly sampled white noise series $y \sim \mathcal{N}(0,1)$ of size $N_s + 1$ with a cadence $\Delta t$ and a total duration $N_s \Delta t$. Then, the time series $y$ is filtered using a time-dependent filter $\hat{h}(f,t)$. The filtering is performed in the frequency domain. For a colored noise, we define a filter in Fourier domain $\hat{h}(f,t)$ at time $t$ as

$$\hat{h}(f,t) = \begin{cases} 0 & \text{for } f < 1/N_s\Delta t \\ g(f,t) H_0 \left(\frac{f}{f_{yr}}\right)^{-\alpha_0/2} & \text{for } f \geq 1/N_s\Delta t, \end{cases} \tag{3.4}$$

where $H_0$ is the gain of the filter, and $\alpha_0$ is the slope. The gain and slope can be controlled over time using the function $g(f,t)$.

For a time series $y$ and its Fourier transform $\hat{y}(f)$, the output $y'(t)$ of the filter is

$$y'(t) = \mathcal{F}^{-1}(\hat{h}(f,t)\hat{y}(f)), \tag{3.5}$$

with $\mathcal{F}^{-1}$ the inverse Fourier transform.

Finally, we fit a quadratic polynomial of the form $a + bt + ct^2$ to the generated time series. The obtained polynomial is then subtracted from the time series. This final step mimics the timing model fit performed in PTA data [7,8], which accommodates the quadratic spin-down (deceleration) of the pulsar rotation while also eliminating part of the moving average component attributed to the nonstationarity.

### C. Setting the model and priors

To test the model, we aim to estimate the Bayes factor for the nonstationary case versus the stationary case given the generated data. As mentioned earlier, we want to restrict ourselves to first-order variations in time of the power-law spectrum amplitude and spectral index. We evaluate the covariance matrix in Eq. (2.6) on a discrete set of frequencies $f_i = i/\tau$ with $i = [1, \ldots, N_f]$ an integer, $N_f$ the size of the basis, and $\tau$ the considered stationarity timescale. In this context, the function $g(f_i,t)$ can be expressed as

$$\log_{10} g(f_i,t) = a\frac{(t-t_0)}{T} - \frac{1}{2} b \frac{(t-t_0)}{T} \log_{10}(f_i/f_{yr}), \tag{3.6}$$

where $T$ is the total time of observation, $f_{yr} = 1/(1\ \text{yr})$ is a reference frequency, and $t_0$ is the initial time.

In this work, we chose to fix $t_0$ at the start time of the entire dataset to directly measure the evolution of amplitude and spectral index from starting values $\log_{10} A(t_0)$ and $\gamma(t_0)$.

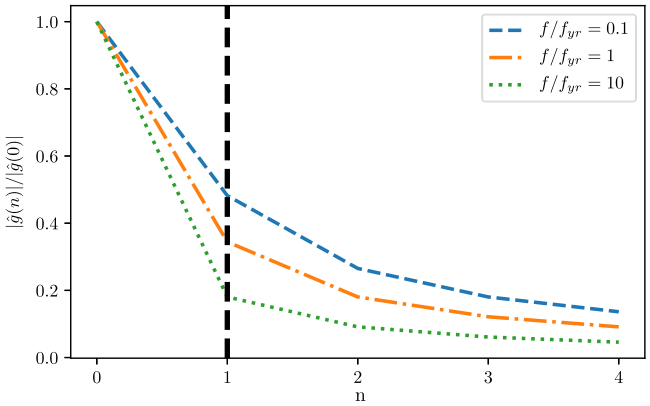


FIG. 2. Norm of the Fourier transform of $g(f_i,t)$ expressed in Eq. (3.6) for $a = 1$, $b = 1$, and three ratios $f_i/f_{yr}$. The $x$ axis corresponds to the harmonics of $1/T$. We show $|\hat{g}(n)|/|\hat{g}(0)|$ for better visibility and comparison.

To perform inference of the parameters, we need to set prior probability distributions. For a slowly evolving signal observed within a time interval $T$, we assume that the stationarity timescale $\tau \simeq T$, so the chosen function $g(f_i,t)$ must have most of its Fourier spectrum concentrated around the lowest frequency $1/T$. This ensures that the 2D spectral measure is concentrated on the diagonal $f = f'$, minimizing leakage between different frequencies (see Appendix A 1). If not, $g(f_i,t)$ is rapidly evolving, and the assumption of local stationarity is no longer valid.

In Fig. 2, we display the Fourier spectrum of $g(f_i,t)$ on the interval $[t_0, t_0 + T]$, denoted as $\hat{g}(n)$, where $n$ corresponds to the harmonics of $1/T$. We can observe that, for parameter values $a = b = 1$, the condition is roughly fulfilled. However, as we decrease the ratio $f_i/f_{yr}$ in the expression of $g(f_i,t)$, the condition is less respected. This implies that the Fourier transform $\hat{g}(n)$ of the function $g(f_i,t)$ has a wider spread at lower $f_i$, leading to more significant changes occurring within timescales shorter than $T$. Thus, we choose $a = b = 1$ to represent typical parameter values where the assumption of a slowly evolving spectrum begins to break down for $T \sim 10$ yr. We decide to set a normal distributed prior $\mathcal{N}(0,1)$ for $a$ and $b$ (see Table I) to confine their values around the interval $[-1,1]$. The inference of parameters may not be precise enough to accurately estimate the evolutionary PSD of the process within the Bayesian scheme (resulting in wide posterior uncertainty). In this sense, this model remains an approximation. Nevertheless, we should be able to capture the nonstationarity of the process through the evaluation of the Bayes factor if $a$ and $b$ deviate too much from 0. This last point is crucial to test whether we are actually dealing with nonstationary noise.

TABLE I. Prior probability distributions used for the parameters of a nonstationary colored noise.

| Parameter | Prior probability |
|---|---|
| $\log_{10} A(t_0)$ | Uniform$(-18, -10)$ |
| $\gamma(t_0)$ | Uniform$(1, 6)$ |
| $a$ | $\mathcal{N}(0, 1)$ |
| $b$ | $\mathcal{N}(0, 1)$ |

### D. Simulation results

Following the procedure presented in Sec. III B, we generate 1000 simulations with 500 samples and 20 years of data containing nonstationary red noise with a low white noise level of $10^{-9}$ s. For each simulation, the parameters $a$ and $b$ are drawn from their respective prior probability distribution, while the initial gain $H_0$ and slope $\alpha_0$ are fixed at $10^{-7}$ and 3, respectively. Inference is performed using the expressions of the likelihood and posterior probability presented in Sec. III A. The prior distributions we use are given in Table I. We employ nested sampling to explore the parameter space for both stationary and nonstationary models to estimate the Bayes factors [25]. We denote the Bayes factor for nonstationarity against stationarity of noise as $\mathcal{B}_S^{\mathrm{NS}}$.

Figure 3 illustrates the precision of nonstationary parameter recovery. We present a histogram showing the differences in each simulation between the injected and recovered maximum likelihood values of parameters $a$ and $b$. The uncertainties are narrower than the prior distribution, indicating that the model is informative. However, the histogram for $b$ appears broader than parameter $a$, which makes an accurate estimate of the evolution of the spectral index challenging.

Figure 4 demonstrates the ability to accurately predict nonstationarity through the evaluation of the Bayes factor $\mathcal{B}_S^{\mathrm{NS}}$. We present a histogram of the Bayes factors evaluated for each simulation. The distribution peaks at low Bayes factors, which corresponds to the fact that we draw $a$ and $b$ from $\mathcal{N}(0, 1)$, hence they are often near zero. If $a$ and $b$ are small, the stationary model should be equally preferred. Additionally, we observe that the Bayes factor rapidly increases if $a$ and $b$ significantly deviate from zero. However, in this simulation, due to the large uncertainties on the recovery of $b$, the information it provides is limited, and the recovered maximum likelihood value does not effectively convey this lack of information (see Appendix A 2). In Fig. 4, the nonstationarity level $\sqrt{a^2 + b^2}$ and the increase in the Bayes factor are predominantly influenced by the value of parameter $a$.

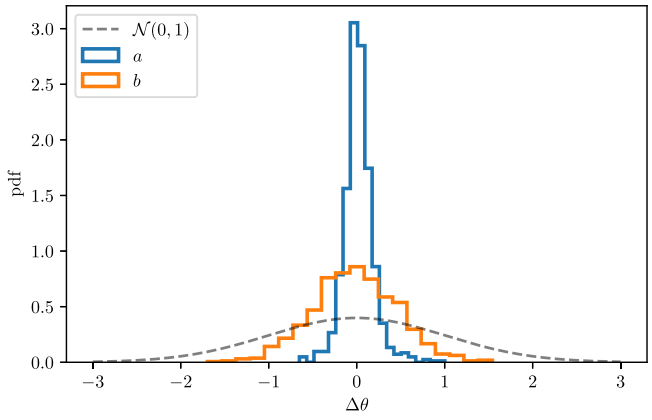


FIG. 3. Normalized histogram of the differences $\Delta\theta$ between injected and recovered maximum likelihood parameter value.

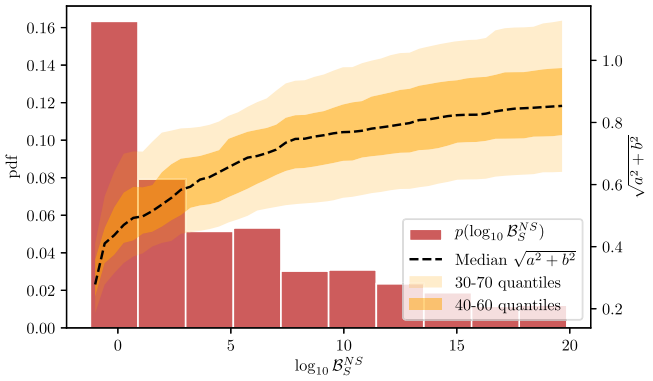


FIG. 4. Normalized histogram of the Bayes factor $\mathcal{B}_S^{\mathrm{NS}}$ nonstationary versus stationary noise obtained from each simulation. We define the level of nonstationarity $\sqrt{a^2 + b^2}$ combining parameters $a$ and $b$. We plot the distribution obtained from their recovered maximum likelihood values for the corresponding simulation and Bayes factor in the $x$ axis.

## IV. RESULTS

The results section will focus on the search for a GWB in the second data release of the EPTA Collaboration [7]. The following analyses use the same methods, priors, and noise models presented in [6,8]. We use the 24.7 years of dataset referred to as DR2full in [6]. The search is for a GWB with a power-law spectrum modeled as stochastic noise spatially correlated between pairs of pulsars, following the so-called Hellings-Downs correlation pattern [3]. The individual pulsar noise models are constructed from a combination of RN and DM noise, also modeled with power-law spectra [8]. These pulsar noise models are built using ENTERPRISE [26].

First, we study the GWB using the stationary formulation of the covariance matrix, as described in [6], but across different time slices of the dataset. Then, we investigate the GWB using the nonstationary formulation of the covariance that was developed in the previous sections.

### A. Time-slice analysis of GWB

Time-slice analyses of the GWB were previously conducted by the Parkes Pulsar Timing Array Collaboration for their third data release [5], where the dataset was segmented

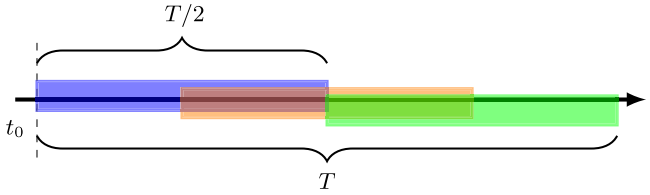


FIG. 5. Timeline representation (to scale) of the three analyzed slices represented by the colors blue, orange, and green with respect to the full time of observation $T$.

into several slices and analyzed as a stationary GWB in each slice. In this section, we perform a similar analysis on three different slices of our dataset. For $T$, the total time span of observation, and $t_0$, the initial time, we analyze the slices $[t_0, T/2]$, $[T/4, 3T/4]$, and $[T/2, T]$, as illustrated in Fig. 5.

In all three slices, we search for a stationary GWB with a power-law spectrum on a 30-frequency linearly spaced Fourier basis with the lowest frequency $2/T$, allowing the spectral index $\gamma$ and amplitude $\log_{10} A$ to vary (see Ref. [6] for details). We denote the recovered GWB posteriors in slices 1, 2, and 3 as $\mathrm{GWB}_1$, $\mathrm{GWB}_2$, and $\mathrm{GWB}_3$ respectively. The results are presented in Fig. 6.

The $\mathrm{GWB}_1$ posterior appears to be much broader than the others. This is expected because the cadence of early observations was lower than those of today (some pulsars may have a cadence of 1 point/month in the early data versus one every 3 days in the latest). The level of white noise is directly influenced by the cadence of observation.[1] The interplay between the power-law spectrum of red noise and the flat white noise tail may strongly influence the estimate of the red noise parameters. Moreover, the quality of the backends' processing raw data from the radio observatories has improved over time. Their ability to simultaneously observe incoming photons at multiple frequency bands (larger bandwidth) allows for a better characterization of the DM noise due to pulses traveling through the interstellar medium [8]. The latter induces a delay in the time of arrival of the pulses with an amplitude that is inversely proportional to the square of the incoming photon's frequency. Its characterization requires good multiband observing capabilities; otherwise, it could be mistaken for RN and contaminate GWB parameter estimation. Improved cadence and noise characterization should provide better constraints on the GWB posteriors, which could explain why $\mathrm{GWB}_1$ appears less informative than $\mathrm{GWB}_2$ and $\mathrm{GWB}_3$. Moreover, if some of the individual pulsar noises are nonstationary, they may leak into the GWB signal. This last point will be explored in the next subsection. It is important to note that this first slice,

[1]For an evenly sampled dataset with measurement uncertainties $\sigma$ and red noise, the PSD can be expressed as $S(f) = Af^{-\gamma} + 2\sigma^2\Delta t$, where $\Delta t$ is the cadence of observation that directly influences the tail of the spectrum [27].

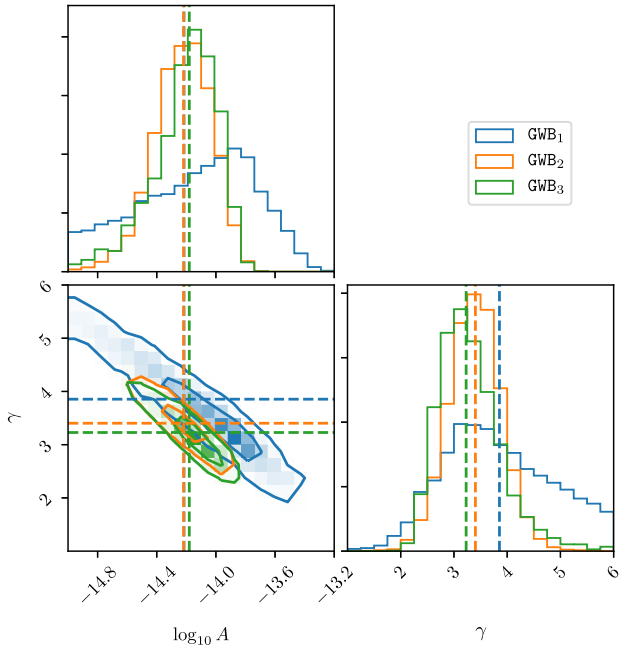


FIG. 6. 2D posterior distributions of GWB parameters analyzed in three time slices. The contours on the 2D histograms are the 68% and 95% credible regions.

$\mathrm{GWB}_1$, corresponds to the first data release from EPTA, where no GWB was less significant. The recovered posteriors are similar to what was observed in [28].

The $\mathrm{GWB}_2$ and $\mathrm{GWB}_3$ posteriors nearly overlap and closely resemble the recovered posteriors using the latest ten years of data in [6]. We observe a slight shift of the spectral index $\gamma$ toward lower values. Figure 7 displays the 1D posteriors of $\gamma$ for the three slices, indicating an estimated decrease in $\gamma$ of approximately 0.66 from the median values. Regarding the amplitude, the median values show little to no evolution. However, evaluating the evolution of $\gamma$ and $\log_{10} A$, especially for $\mathrm{GWB}_1$, is challenging due to the large uncertainties.

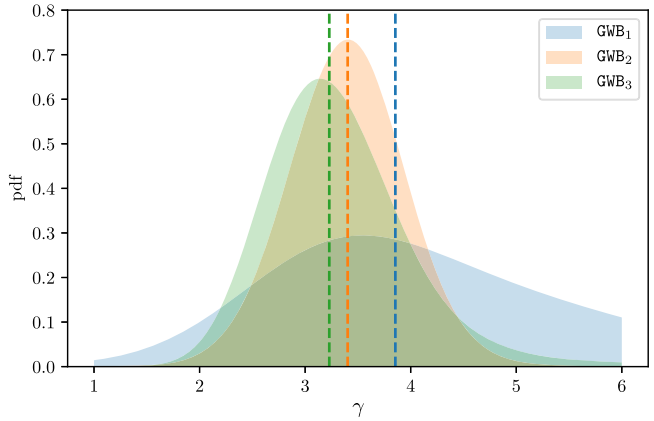


FIG. 7. 1D marginalized posterior distribution of GWB spectral index $\gamma$ analyzed in three time slices.

One caveat of this analysis is that we lose information about the low-frequency range after performing the slices. Indeed, the lowest sampled frequency bin is at frequency $1/T$, and increasing $T$ raises its value. In PTAs, the lowest frequencies contain important information when searching for a GWB. Additionally, this analysis prevents us from estimating a Bayes factor in favor of nonstationary behavior against stationarity. The model presented in this paper should address both issues.

### B. Nonstationary power-law spectrum GWB

We will now analyze the data using a GWB model with HD correlations and an evolutionary power-law spectrum. In Sec. II B, we developed an expression for such a model.

The previous time-slice analysis only reveals small variations in the amplitude and spectral index of the power law. For this reason, we chose to limit ourselves to the simplest possible model where the evolution of amplitude and spectral index is approximated to first order in time. We have

$$\log_{10}A(t) \sim \log_{10}A(t_0) + a(t-t_0)/T,$$
$$\gamma(t) \sim \gamma(t_0) + b(t-t_0)/T, \tag{4.1}$$

where $t_0$ is the minimum time of the dataset, $a$ is the parameter controlling the evolution of log-amplitude, and $b$ is the parameter controlling the evolution in $\gamma$.

We performed the analysis using a normal prior $\mathcal{N}(0,1)$ on $a$ and $b$ and show the recovered posteriors in Fig. 8. We find a GWB with $\log_{10}A(t_0) = -14.4^{+0.2}_{-0.2}$, $\gamma(t_0) = 3.7^{+0.5}_{-0.5}$, $a = 0.0^{+0.3}_{-0.3}$, and $b = -0.7^{+0.7}_{-0.6}$, fitting a spectrum with a constant amplitude and decreasing spectral index. This behavior aligns with the estimated evolution of the parameters using the medians of the 1D posteriors in the slice analysis. Still, the uncertainties on the estimate of parameter $b$ are very large and $b = 0$ lies within the 1-$\sigma$ credible interval. For this GWB model, we estimate a Bayes factor of nonstationarity against stationarity $\mathcal{B}^{\rm NS}_S = 1.5$, indicating no evidence for a nonstationary GWB. The apparent evolution of the spectral properties of this process is likely due to improvements in receiver-backend systems and the enhanced cadence of observation over the years. As demonstrated in the previous section, the inferred posteriors of the GWB parameters in the oldest slice of our dataset are very uninformative. This limitation may impede our ability to characterize the time-dependent features of the spectrum effectively.

### C. Nonstationary individual noise

Earlier we mentioned the possible nonstationarity of individual pulsar noise. In the Introduction, we presented the two dominating sources of time-correlated noise in pulsar timing: the RN, due to stochastic variation of the pulsar's spin rate, and the DM noise, due to the variation of the interstellar medium electron density throughout the travel of the pulses to Earth [8]. Another source of frequency-dependent noise is identified, chromatic noise, with an amplitude proportional to $\nu^{-4}$ with $\nu$ the incoming photon frequency. This type of noise is assumed to be stationary in PTA data analysis. Incorrect modeling of pulsar noise could deteriorate the inference of the GWB parameters.

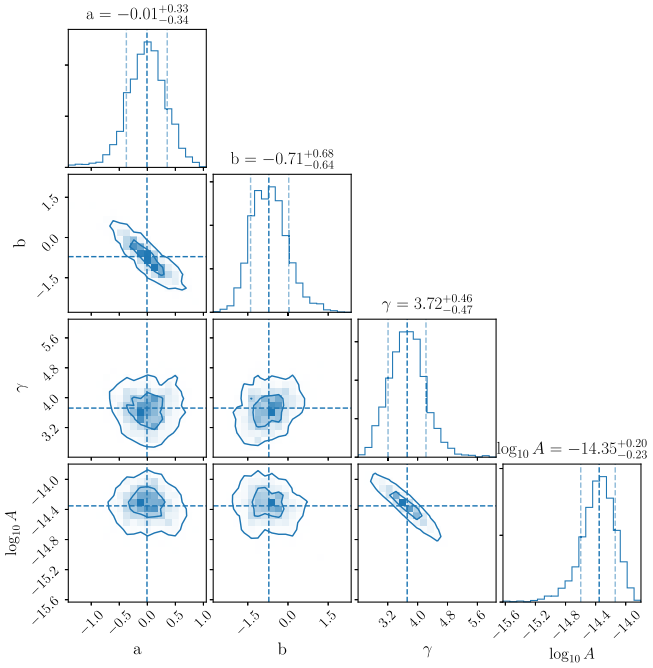


FIG. 8. Corner plot of nonstationary GWB parameters inferred from the EPTA dataset. The contours on the 2D histograms are the 68% and 95% credible regions. The dashed lines on the 1D histograms show the 14%/50%/86% percentiles.

In this section, we compare the stationary formulation of the custom noise models presented in [8] to the nonstationary formula of Sec. II B in order to assess the stationarity of pulsar noises. Every noise spectrum is modeled as a power-law spectrum. We only test the first-order model with function $g(f,t)$ given by Eq. (3.6) to allow each noise component to vary with time. We perform MCMC analysis using product-space sampling [29] to estimate the Bayes factor $\mathcal{B}^{\rm NS}_S$. Results are gathered in Table II.

According to the distribution of Bayes factors, it appears that the pulsar noise does not exhibit first-order variations in time, with the exception of J1713 + 0747 for which $\log_{10}\mathcal{B}^{\rm NS}_S = 4.3$, hence favoring the nonstationary model.

Given the noise parameter values in Table III, it seems that the recovered nonstationarity originates from the DM noise with a relatively well constrained $a_{\rm DM} - 0.7^{+0.2}_{-0.2}$ indicating that the DM amplitude decreases with time and $b_{\rm DM} = 0.4^{+0.6}_{-0.6}$ indicating an increasing spectral index (though with large posterior uncertainties). For the RN, the $a_{\rm RN}$ and $b_{\rm RN}$ parameters are consistent with zero within the 1-$\sigma$ credible interval suggesting stationary behavior.

TABLE II. Bayes factors for nonstationary against stationary noise models. The pulsar with $\log_{10}\mathcal{B}_S^{\rm NS} > 2$ is shown in bold font. We observe that, in general, pulsars do not exhibit evidence of nonstationary noise, except for J1713 + 0747, for which we show the two values of $\log_{10}\mathcal{B}_S^{\rm NS}$ obtained with (left) and without (right) including the exponential dip events in the model. The case of J1713 + 0747 is discussed in the text.

| Pulsar | $\log_{10}\mathcal{B}_S^{\rm NS}$ |
|---|---|
| J0030 + 0451 | 0.0 |
| J0613 − 0200 | −0.6 |
| J0751 + 1807 | −0.3 |
| J0900 − 3144 | −0.3 |
| J1012 + 5307 | −1.2 |
| J1022 + 1001 | −0.2 |
| J1024 − 0719 | −0.3 |
| J1455 − 3330 | 0.0 |
| J1600 − 3053 | 0.0 |
| J1640 + 2224 | −0.2 |
| **J1713 + 0747** | **4.3**/−0.6 |
| J1730 − 2304 | 1.1 |
| J1738 + 0333 | 0.1 |
| J1744 − 1134 | −0.5 |
| J1751 − 2857 | −0.2 |
| J1801 − 1417 | 0.3 |
| J1804 − 2717 | 0.2 |
| J1843 − 1113 | −0.4 |
| J1857 + 0943 | −0.4 |
| J1909 − 3744 | 0.7 |
| J1910 + 1256 | 0.0 |
| J1911 + 1347 | −0.4 |
| J1918 − 0642 | −0.3 |
| J2124 − 3358 | −0.3 |
| J2322 + 2057 | 0.0 |

However, J1713 + 0747 is known to have challenging noise properties [30–32] and nonstationary DM events were already reported for this pulsar in [33]. In [8], two exponential dip events were reported in the timing residuals of J1713 + 0747. An exponential dip is a deterministic DM-like effect that could be induced by extreme scattering events in the interstellar medium or a sudden change in the pulse profile of the pulsar, introducing transitory frequency-dependent timing delays that can last up to a few months [34,35]. In Table III, we see that the inclusion of the exponential dips in our model removes the apparent nonstationary behavior of J1713 + 0747, now showing $a_{\rm DM} = -0.2^{+0.3}_{-0.2}$ and $b_{\rm DM} = -0.4^{+0.7}_{-0.7}$. The Bayes factor for nonstationary noise versus stationary noise after the inclusion of exponential dip events is now $\log_{10}\mathcal{B}_S^{\rm NS} = -0.6$.

Although not the best description, the nonstationary model was able to hint toward the presence of time-dependent features in the timing residuals of pulsar J1713 + 0747 through the evaluation of the Bayes factor. We tried searching again for a GWB in the EPTA dataset, now considering the nonstationarity (exponential dips) of J1713 + 0747, and found no difference in the estimate of the Bayes factor in favor of HD or in the inferred GWB parameters with respect to [6]. The stationary model of individual pulsar noise seems to be enough an approximation when searching for a GWB.

TABLE III. Table of median noise parameters (RN and DM) inferred from pulsar J1713 + 0747 timing data for the nonstationary model (NS), nonstationary model with exponential dips (NSed), and the stationary model with exponential dips (Sed). The displayed quantiles are 0.14 and 0.86.

| Parameter | Median (NS) | Median (NSed) | Median (Sed) |
|---|---|---|---|
| $\gamma_{\rm DM}$ | $1.0^{+0.4}_{-0.4}$ | $1.9^{+0.6}_{-0.6}$ | $1.6^{+0.2}_{-0.2}$ |
| $\log_{10} A_{\rm DM}$ | $-13.0^{+0.1}_{-0.1}$ | $-13.6^{+0.2}_{-0.2}$ | $-13.7^{+0.04}_{-0.04}$ |
| $a_{\rm DM}$ | $-0.7^{+0.2}_{-0.2}$ | $-0.2^{+0.3}_{-0.2}$ | ⋯ |
| $b_{\rm DM}$ | $0.4^{+0.6}_{-0.6}$ | $-0.4^{+0.7}_{-0.7}$ | ⋯ |
| $\gamma_{\rm RN}$ | $3.2^{+0.8}_{-0.8}$ | $3.2^{+0.8}_{-0.8}$ | $3.5^{+0.6}_{-0.6}$ |
| $\log_{10} A_{\rm RN}$ | $-13.9^{+0.4}_{-0.5}$ | $-14.0^{+0.4}_{-0.5}$ | $-14.3^{+0.3}_{-0.3}$ |
| $a_{\rm RN}$ | $-0.3^{+0.5}_{-0.5}$ | $-0.2^{+0.5}_{-0.5}$ | ⋯ |
| $b_{\rm RN}$ | $-0.2^{+0.9}_{-0.9}$ | $-0.3^{+0.9}_{-0.9}$ | ⋯ |

### D. Eccentric single source as nonstationary GWB

If a binary system of SMBHs is eccentric, it will radiate polychromatic GWs, contrary to the monochromatic circular binaries [36]. On average, the GW signal can be decomposed as a sum of frequencies, which are the harmonics of the orbital frequency of the binary system. The waveform of an eccentric GW shows that the instantaneous frequency of the signal changes significantly within one orbital period [9,10].

In [37], it was shown that single binaries could be mistaken for a GWB with a power-law spectrum. For eccentric binaries with orbital periods comparable to the total time of observation of a PTA, the instantaneous change in frequency of the GW signal can be seen as a GWB with evolving spectral index $\gamma(t)$ (see Fig. 9). Indeed, the $\gamma$ controls the balance between low and high frequencies of the power-law spectrum and the presence of an eccentric signal would introduce time-dependent features. GW sources of higher eccentricity and higher amplitude would introduce more nonstationarity because of their stronger frequency evolution or higher signal to noise ratio (SNR). In this section, we want to investigate what we would see when a stationary GWB is influenced by an eccentric single source with low SNR.

We create a simulation with 10 pulsars (isotropically distributed in the sky), 15 years of observation, and only white noise at a level of $10^{-7}$ s. The simulation contains a stationary GWB with $\log_{10} A = -14.2$ and $\gamma = 3$. We generate two cases: (i) only a stationary GWB is present in the data; (ii) the same stationary GWB is accompanied by an eccentric source located at the Virgo galaxy cluster with amplitude $h = 10^{-13.5}$ and orbital period of 15 years

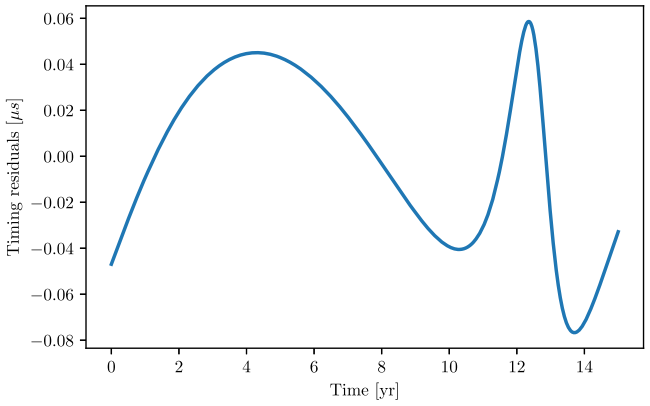


FIG. 9. One period of an eccentric signal (Earth term only) with eccentricity 0.6 and orbital period of 15 years. We see that the instantaneous frequency significantly changes within one period. The waveform was generated using ENTERPRISE extensions [38].

(here corresponding to SNR $\simeq 1.8$). We produce eight datasets with different eccentricities of the source $e$ ranging from 0.1 to 0.8. We analyzed these datasets with the nonstationary model presented in Sec. II B, considering HD correlations. In Fig. 10, we show that the presence of an eccentric source with $e = 0.5$ introduces nonstationarity in

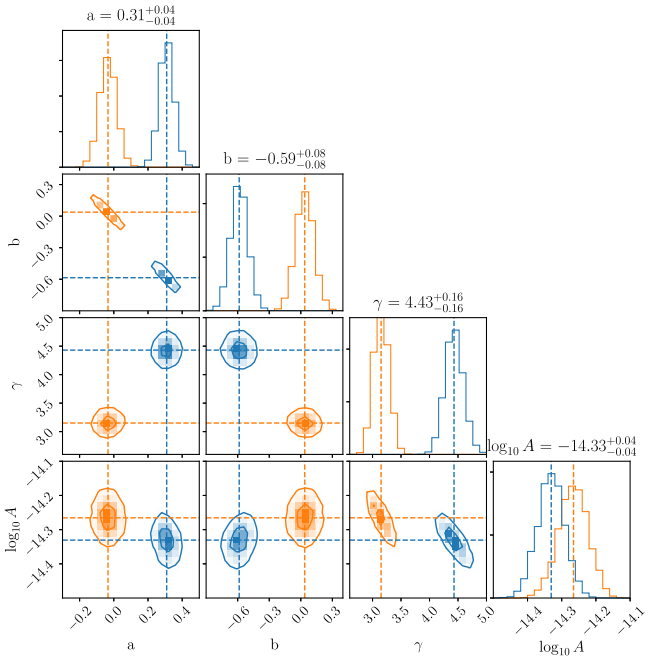


FIG. 10. Corner plot of nonstationary GWB parameters. In orange, we show the results for a simulated stationary GWB with $\log_{10} A = -14.2$ and $\gamma = 3$. In blue, we show the results for the same realization of GWB "contaminated" by an eccentric single source with 15 years orbital period and eccentricity $e_0 = 0.5$. The dashed lines are the medians. The displayed values of parameters correspond to the medians of the orange histograms. The contours on the 2D histograms are the 68% and 95% credible regions.

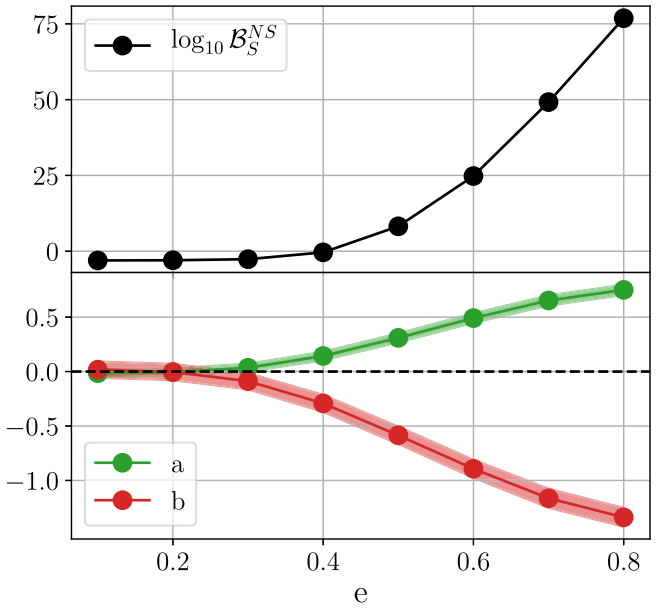


FIG. 11. Top: Bayes factor for nonstationary against stationary GWB models as a function of the eccentricity of the injected source. Bottom: recovered posterior median of nonstationary parameters $a$ and $b$ as a function of eccentricity. The shaded region around $a$ and $b$ represents the 1-$\sigma$ credible interval.

the inferred GWB parameters. For case (i), parameters $a$ and $b$ are near zero, meaning that we are dealing with a stationary signal. In case (ii), where the eccentric signal is introduced, they deviate significantly from zero. Note that, in our simulation, the parameter $b$ is negative (i.e., $\gamma$ is decreasing) showing that some specific realizations of eccentric binaries can somewhat reproduce the behavior observed in the real data.

The evolution of the recovered Bayes factor $\mathcal{B}_S^{\mathrm{NS}}$ and parameters $a$ and $b$, as a function of the eccentricity $e$ of the injected GW signal, is illustrated in Fig. 11. It is evident that the model effectively captures the nonstationary features induced by the presence of the eccentric source when the latter is loud enough. The Bayes factor begins to favor the nonstationary model for eccentricities higher than 0.4. For this analysis, we have utilized the expression of $g(f,t)$ from Eq. (3.6). In future work, it might be beneficial to design a $g(f,t)$ that better suits this problem.

## V. DISCUSSION AND CONCLUSION

The derived kernel serves the purpose of modeling nonstationarity in noise. It resembles the stationary kernels commonly used in PTA analyses, but incorporates the capability to capture the time-dependent variance of the Gaussian weights. In its current form, this kernel remains an approximation and is valid only for small deviations from a stationary model. A significant challenge lies in understanding how correlations between frequencies can be accurately modeled, leading to an

exact and nonapproximated version of the nonstationary kernel. Further exploration in this area is necessary to improve our understanding and refine the methodology.

This approach necessitates a prior assumption about the expression of the evolutionary PSD. One significant advantage is its consistent expression of the covariance matrix for the entire dataset, unlike the time-slice analysis, which necessitates cutting the data into several segments. This can be particularly valuable for applications like PTAs, where noise is most pronounced in the lowest frequency bins.

In Sec. III D, we showed that our model can occasionally produce broad posterior uncertainties for the nonstationary parameters, and the stationary solution may fall within the 1-$\sigma$ credible interval. Consequently, accurately characterizing the evolutionary PSD can be challenging. Thus, it becomes crucial to assess the Bayes factor against the stationary model to determine whether the preference lies with a genuinely nonstationary signal.

The slice analysis reveals a clear evolution in the inferred GWB properties. The first epoch appears much less informative than the following two. This can be attributed to an evolution of the data quality over time or the leakage of nonstationary features of individual pulsar noises into the GWB estimate. We conducted individual analyses for each pulsar to investigate nonstationary features and found that only J1713 + 0747 seemed to exhibit time-dependent DM noise amplitude variations. Nonstationary DM events were already reported for this pulsar in previous studies. Exponential dip events offer the best description of these events.

The nonstationary model developed indicates that there is no evidence for a nonstationary GWB in the second data release of the EPTA data. Based on the results of the slice analysis, we infer that the apparent evolution of the inferred parameter values may stem from a poor characterization of individual pulsar noise contaminating the observed GWB. However, there is still a possibility that the presence of GW signals induced by eccentric SMBHBs in the data could produce similar nonstationary behaviors. The last section explored this hypothesis and showed that PTAs could detect nonstationary features of the GWB if an eccentric source is present in the data.

With the current dataset, it remains inconclusive whether the observed evolution solely results from a change in backend quality or represents a genuine nonstationary feature of the GWB. Furthermore, if the noise from individual pulsars exhibits nonstationarity, it could potentially contaminate the GWB estimate, making it appear as nonstationary. Further investigation into this matter is warranted in future studies. Additionally, the upcoming release of the International Pulsar Timing Array third data release, with enhanced observation cadence and improved frequency coverage, is expected to provide better noise characterization and definitive insights into this issue.

The EPTA data are publicly available, see Ref. [39].

## ACKNOWLEDGMENTS

M. F. thanks Stas Babak, Philippa Cole, Aurélien Chalumeau, and Hippolyte Quelquejay for the feedback on the manuscript. M. F. also thanks Philippe Bacon and Quentin Baghi for the interesting discussions. M. F. acknowledges financial support from MUR PRIN, Grant No. 2022-MYL2X, project "GRAPE," funded by the European Union—Next Generation EU. The Nançay Radio Observatory is operated by the Paris Observatory, associated with the French Centre National de la Recherche Scientifique (CNRS), and partially supported by the Region Centre in France. I. C., L. G., and G. T. acknowledge financial support from "Programme National de Cosmologie and Galaxies" (PNCG) and "Programme National Hautes Energies" (PNHE) funded by CNRS/INSU-IN2P3-INP, CEA, and CNES, France. I. C., L. G., and G. T. acknowledge financial support from Agence Nationale de la Recherche (ANR-18-CE31-0015), France. The Sardinia Radio Telescope (SRT) is funded by the Department of University and Research (MIUR), the Italian Space Agency (ASI), and the Autonomous Region of Sardinia (RAS) and is operated as a National Facility by the National Institute for Astrophysics (INAF). D. P. acknowledges support from the INAF Large Grant "Gravitational Wave Detection using Pulsar Timing Arrays" (P. I. D. Perrodin, 2023). J. A. acknowledges support from the European Commission (ARGOS-CDS; Grant Agreement No. 101094354). G. M. S. acknowledges financial support provided under the European Union's H2020 ERC Consolidator Grant "Binary Massive Black Hole Astrophysics" (B Massive, Grant Agreement No. 818691). The Westerbork Synthesis Radio Telescope is operated by the Netherlands Institute for Radio Astronomy (ASTRON) with support from the Netherlands Foundation for Scientific Research (NWO).

## APPENDIX: TO BE PROVIDED BY AUTHOR

### 1. 2D spectral measure

The generalized Wiener-Khintchine relations are defined in [18] as

$$C(t,t') = \int df df' S(f,f') e^{2\pi i (tf - t'f')},$$
$$S(f,f') = \int dt dt' C(t,t') e^{-2\pi i (tf - t'f')}. \tag{A1}$$

From the definition of the nonstationary covariance matrix in Eq. (2.6), we can estimate the 2D spectral density,

$$S(f,f') = \int dt dt' \sum_i \sigma_i^2 g_i(t) g_i(t') \left\{ e^{-2\pi i t (f+f_i)} e^{-2\pi i t' (-f-f_i)} + e^{-2\pi i t (f_i - f)} e^{-2\pi i t' (f' - f_i)} \right\} \tag{A2}$$

with $\hat{g}_i(f)$ the Fourier transform of the function $g_i(t)$, we have

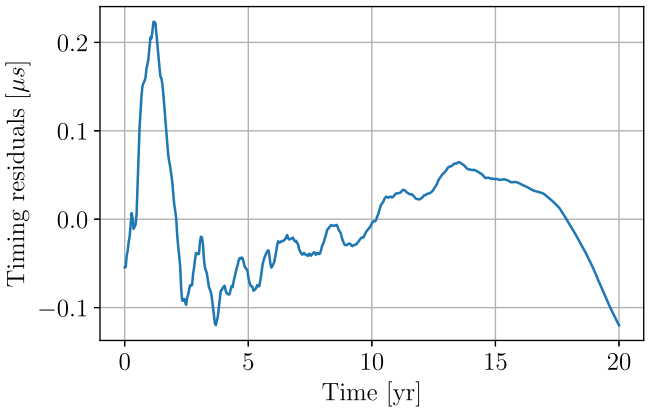


FIG. 12. Time-domain plot of a nonstationary noise generated with the method presented in Sec. III B.

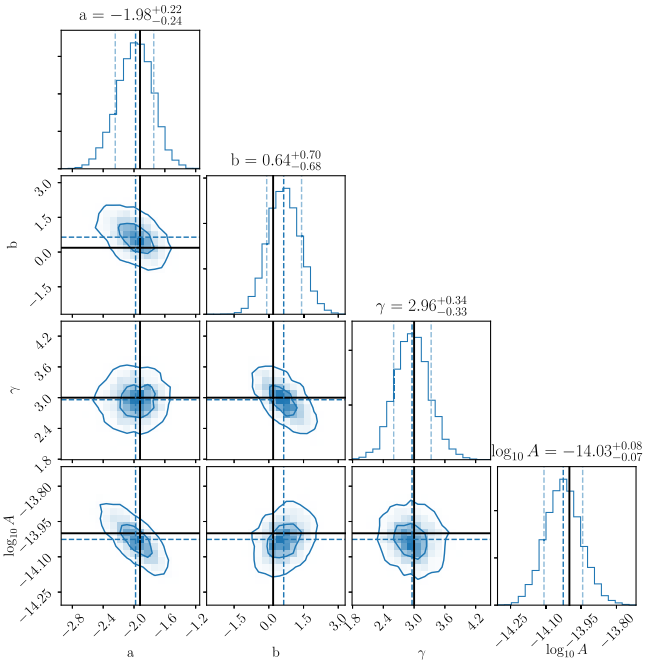


FIG. 13. Corner plot of nonstationary parameters inferred from the simulation in Fig. 12. The solid black lines show the injected values of parameters. The contours on the 2D histograms are the 68% and 95% credible regions. The dashed lines on the 1D histograms show the 14%/50%/86% percentiles.

$$S(f,f') = \sum_i \frac{\sigma_i^2}{2}[\hat{g}_i(f+f_i)\hat{g}_i(-f'-f_i) + \hat{g}_i(-f+f_i)\hat{g}_i(f'-f_i)]. \quad \text{(A3)}$$

Note that if $\hat{g}_i(f) = \delta_i(f)$ (a Dirac delta) and applying Eq. (A1), we recover the stationary covariance matrix of Eq. (2.3), giving a 2D spectral measure $S(f,f')$ that is only defined on the diagonal $f = f'$ and evaluated at discrete frequencies $f_i$. The spread of the function $\hat{g}_i(f)$ determines how much we deviate from the stationary covariance matrix.

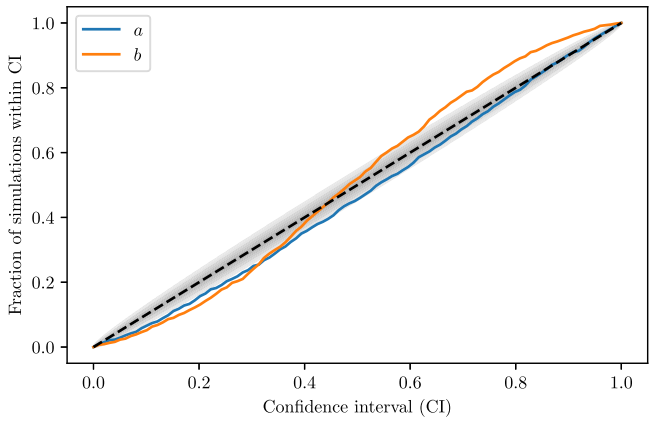


FIG. 14. PP plot comparing the fraction of simulations within confidence interval (CI) and CI. The gray shaded region represents the 1-$\sigma$ error from $y = x$.

### 2. Simulations

In Sec. III B, we presented a method to generate a nonstationary colored noise. Here, we show an example of one realization of nonstationary noise using this method. In Fig. 12, we show an example of one realization of nonstationary noise using this method. Fig. 13 shows the corresponding corner plots with the injected versus recovered values of parameter. We can see that the model performs well, except for parameter $b$, where the posterior uncertainties are quite large.

### 3. PP plots

In [40], a validation method is presented. It necessitates both data-generation and model-fitting software. In this context, our model-fitting software is PTMCMC [24], which has been previously utilized. During the fitting process, we sample the posterior distribution employing our model, specifically the nonstationary kernel, using the fitting software. In each simulation, we possess knowledge of the injected parameter values $\vec{\theta}_0$, allowing us to calculate the corresponding quantiles $q(\vec{\theta}_0)$ with respect to the posterior distribution derived from the fitting software. The methodology in [40] demonstrates that if the fitting software operates accurately and the generated data adhere to the same posterior distribution for parameters as the constructed model, then the quantiles $q(\vec{\theta}_0)$ must be distributed as Uniform(0, 1).

In Fig. 14 we show the PP plot for the quantiles $q(\vec{\theta}_0)$ obtained from simulations. They seem to follow the relationship $y = x$, but we notice discrepancies, especially for parameter $b$ which is the one with larger uncertainty.

[1] A. Sesana, A. Vecchio, and C. N. Colacino, Mon. Not. R. Astron. Soc. **390**, 192 (2008).
[2] E. S. Phinney, arXiv:astro-ph/0108028.
[3] B. Allen and J. D. Romano, Phys. Rev. D **108**, 043026 (2023).
[4] G. Agazie *et al.*, Astrophys. J. Lett. **951**, L8 (2023).
[5] D. J. Reardon *et al.*, Astrophys. J. Lett. **951**, L6 (2023).
[6] J. Antoniadis *et al.* (EPTA, InPTA Collaborations), Astron. Astrophys. **678**, A50 (2023).
[7] J. Antoniadis *et al.* (EPTA, InPTA Collaborations), Astron. Astrophys. **678**, A48 (2023).
[8] J. Antoniadis *et al.* (EPTA, InPTA Collaborations), Astron. Astrophys. **678**, A49 (2023).
[9] S. R. Taylor, E. A. Huerta, J. R. Gair, and S. T. McWilliams, Astrophys. J. **817**, 70 (2016).
[10] A. Susobhanan, A. Gopakumar, G. Hobbs, and S. R. Taylor, Phys. Rev. D **101**, 043022 (2020).
[11] J. Ellis and N. Cornish, Phys. Rev. D **93**, 084048 (2016).
[12] N. J. Cornish, Phys. Rev. D **102**, 124038 (2020).
[13] R. van Haasteren and M. Vallisneri, Phys. Rev. D **90**, 104012 (2014).
[14] S. Mallat, G. Papanicolaou, and Z. Zhang, Ann. Stat. **26**, 1 (1998).
[15] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. (McGraw-Hill, Boston, 2002).
[16] C. E. Rasmussen, Gaussian processes in machine learning, in *Advanced Lectures on Machine Learning*, edited by O. Bousquet, U. von Luxburg, and G. Rätsch (Springer, Berlin, Heidelberg, 2004), pp. 63–71.
[17] R. van Haasteren and M. Vallisneri, Mon. Not. R. Astron. Soc. **446**, 1170 (2014).
[18] M. B. Priestley, J. R. Stat. Soc. Ser. B **27**, 204 (1965).
[19] H. Hong, Mech. Syst. Signal Process. **190**, 110131 (2023).
[20] Y.-L. K. Samo and S. Roberts, arXiv:1506.02236.
[21] S. Remes, M. Heinonen, and S. Kaski, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17 (Curran Associates, Red Hook, NY, 2017), pp. 4645–4654.
[22] G. Agazie *et al.*, Astrophys. J. Lett. **951**, L10 (2023).
[23] D. J. Reardon *et al.*, Astrophys. J. Lett. **951**, L7 (2023).
[24] J. Ellis and R. van Haasteren, jellis18/ptmcmcsampler: Official release (2017), 10.5281/zenodo.1037579.
[25] J. Skilling, Bayesian Anal. **1**, 833 (2006).
[26] J. A. Ellis, M. Vallisneri, S. R. Taylor, and P. T. Baker, ENTERPRISE: Enhanced numerical toolbox enabling a robust pulsar inference suite (Zenodo, 2020), 10.5281/zenodo.4059815.
[27] C. J. Moore, S. R. Taylor, and J. R. Gair, Classical Quantum Gravity **32**, 055004 (2015).
[28] L. Lentati *et al.*, Mon. Not. R. Astron. Soc. **453**, 2576 (2015).
[29] S. Hee, W. J. Handley, M. P. Hobson, and A. N. Lasenby, Mon. Not. R. Astron. Soc. **455**, 2461 (2015).
[30] Gabriella Agazie *et al.*, Astrophys. J. Lett. **951**, L10 (2023).
[31] M. T. Lam, Res. Notes AAS **5**, 167 (2021).
[32] J. Singha *et al.*, Mon. Not. R. Astron. Soc. **507**, L57 (2021).
[33] L. Lentati *et al.*, Mon. Not. R. Astron. Soc. **458**, 2161 (2016).
[34] R. M. Shannon *et al.*, Astrophys. J. Lett. **828**, L1 (2016).
[35] W. A. Coles *et al.*, Astrophys. J. **808**, 113 (2015).
[36] P. C. Peters and J. Mathews, Phys. Rev. **131**, 435 (1963).
[37] J. Antoniadis *et al.* (EPTA, InPTA Collaborations), arXiv:2306.16226.
[38] S. R. Taylor, P. T. Baker, J. S. Hazboun, J. Simon, and S. J. Vigeland, ENTERPRISE extensions (2021), v2.4.3, https://github.com/nanograv/enterprise_extensions.
[39] https://epta.pages.in2p3.fr/epta-dr2/; J. Antoniadis *et al.* (EPTA Collaboration), Astron. Astrophys. Rev. **678**, A48 (2023).
[40] S. R. Cook, A. Gelman, and D. B. Rubin, J. Comput. Graph. Stat. **15**, 675 (2006).